DOCUMENT RESUME

ED 236 188 TM 830 695

AUTHOR Deno, Stanley L.; And Others
TITLE The Structure of Instruction Rating Scale (SIRS): Development and Technical Characteristics.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
REPORT NO IRLD-RR-107
PUB DATE Jan 83
CONTRACT 300-80-0622
NOTE 43p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Classroom Environment; Elementary Education; Factor Structure; *Interrater Reliability; *Predictor Variables; *Rating Scales; School Districts; *Test Reliability
IDENTIFIERS *Structure of Instruction Rating Scale

ABSTRACT

Using instructional variables identified by the literature as important in predicting classroom achievement, a bi-polar rating scale was designed to assess the structure of instruction in resource rooms. The data for 158 elementary school children in four school districts were analyzed. The scale evidenced good reliability, both in terms of inter-observer agreement and internal consistency measures. Factor analysis of the scale identified one factor consisting of 9 of the 12 scale variables. Teacher use of instructional structure and implications for educational practice are discussed. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED236188

University of Minnesota

Research Report No. 107

THE STRUCTURE OF INSTRUCTION RATING SCALE (SIRS): DEVELOPMENT AND TECHNICAL CHARACTERISTICS

Stanley L. Deno, Robert King, Russell Skiba, Bonita Sevcik, and Caren Wesson

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM EC
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Yssel Ryke
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IRLD

Institute for Research on Learning Disabilities

TM 830 695

IRLD

Director: James E. Ysseldyke

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Office of Special Education.

Research Report No. 107

THE STRUCTURE OF INSTRUCTION RATING SCALE (SIRS): DEVELOPMENT AND TECHNICAL CHARACTERISTICS

Stanley L. Deno, Robert King, Russell Skiba, Bonita Sevcik, and Caren Wesson

Institute for Research on Learning Disabilities

University of Minnesota

January, 1983

Abstract

Using instructional variables identified by the literature as important in predicting classroom achievement, a bi-polar rating scale was designed to assess the structure of instruction in resource rooms. The data for 158 elementary school children in four school districts were analyzed. The scale evidenced good reliability, both in terms of inter-observer agreement and internal consistency measures. Factor analysis of the scale identified one factor consisting of 9 of the 12 scale variables. Teacher use of instructional structure and implications for educational practice are discussed.

The Structure of Instruction Rating Scale (SIRS):

Development and Technical Characteristics

The construct of classroom structure and its relationship to direct instructional procedures, academic engaged time, and a student's learning rate has emerged as an important variable in research in the field of educational psychology (Fisher, Berliner, Filby, Marliave, Cahen, & Dishaw, 1980; Stevens & Rosenshine, 1981). Stated more broadly, several key variables seem to coalesce or interact that determine effective teaching practices for regular and special education teachers. The concern for delimiting effective teaching practices for special education students has become even more important, since many of these students have a prior history of failure in academic settings. Most notably, the recent introduction of direct instructional practices based on a behavioral model of teaching appears most effective and promising for teachers of special education students and those in inner cities (Cooley & Leinhardt, 1980; Gersten, Carnine, Zoref, & Cronin, 1981; Stevens & Rosenshine, 1981; Zoref, 1981).

The hallmark of a direct instructional model is that it relies on empirically testable and proven experimentally-based instructional procedures; modeling, prompting, shaping, and the use of reinforcement schedules are among the most commonly used techniques selected from the behavioral armamentarium (Becker, Engelmann, & Thomas, 1973). Recently, Starlin (1979) pinpointed four effective teaching practices for reading: (a) demonstrating the desired performance, (b) directly practicing the behavior, (c) delivering positive feedback, and (d) adjusting the curriculum to maxmimize learning. These tactics and

others employed by teachers (e.g., controlled practice that focuses on convergent questions, increased use of direct teacher questions, providing praise for correct answers as well as correction for incorrect answers, and monitoring independent seat work practice) have been shown to be effective instructional tactics that relate significantly to student achievement (Stevens & Rosenshine, 1981). Moreover, the frequency of correct answers has been shown to correlate quite highly with student learning as well as pacing of the instructional lesson (Fisher et al., 1980; Gersten, 1981; Stevens & Rosenshine, 1981).

Consistent with these research findings on student learning is an emphasis on the direct practice of behavior(s) by the students across various subjects such as reading and mathematics (Borg, 1980; Starlin, 1979). Put simply, students must be given the opportunity (time) to practice and engage directly in the learning process. A major finding of a six-year research project funded by the National Institute for Education was that student academic achievement is correlated positively with the amount of relevant engaged time the student spends in reading and mathematics (Borg, 1980; Fisher et al., 1980). Further support for the positive relationship between student achievement and academic engaged time also has been documented by Graden, Thurlow, and Ysseldyke (1982).

Central to the delivery of instruction is the way in which the classroom environment is structured by the teacher. Teachers vary considerably in the value they place on providing instruction on academic tasks. It has been suggested that teachers with "high

academic press" may provide more opportunity for the student to engage in academic tasks (Fisher et al., 1980, p. 11). Miller (1980) described the relationship between the constructs of classroom structure and time: "Deriving from the notion of time is the notion of structure. I view structure as the purposeful ordering or placement of people, materials, and resources in time" (Miller, 1980, p. 163).

Thus, there appears to be evidence that the degree of structure in an instructional lesson provided by a teacher will determine the extent to which a student engages in the practice of an academic task. Moreover, there is evidence that certain teaching practices are correlated highly with student academic achievement. The focus of the current research was to evaluate an observational instrument developed to assess the amount and type of instruction a student received in an instructional setting where teachers were trained in the use of data-based teaching practices.

Method

Development of the Scale

A five-point rating scale that used variables considered necessary for effective instruction was developed. The selection of variables for the instrument was based on current research findings. The variables comprising the Structure of Instruction Rating Scale (SIRS), their operational definitions, and the rating scale format are included in Appendix A. The SIRS, as originally constituted, included only the first 10 variables, and only these variables were observed for the first data collection. The variables Oral Practice and Silent

Practice were added after the first data collection: thus, the SIRS consisted of 12 variables for the second and third data collections.

Two rounds of pilot data were collected, analyzed, and refined to develop the initial 10 variables. It should be noted that the operational definition of the variable Positive Consequences did not include teacher praise. Teacher praise was excluded because of evidence suggesting that the reinforcing value of teacher praise still needs to be empirically validated (Brophy, 1981).

The SIRS purposely was designed to focus on the instruction a student receives, rather than focusing on the teacher. This approach was taken because it is possible, and even likely, that the behaviors of a teacher toward a group of students may in fact be differentially effective with individual students. Since instructional effectiveness is measured by cumulating individual achievement and since special education focuses on the individual, the individual instructional program was selected as the target for measurement rather than the teacher.

Rater Training and Interobserver Agreement

During training, raters observed and rated videotaped recordings of a teacher and two to three students. Two tapes of a teacher, judged by her associates to be highly competent, were made one week apart. In the first session, the teacher was instructed to conduct her class as she normally would. For the second tape, the teacher was asked to model what she considered to be poor teaching techniques. These two tapes provided examples of the extremes represented by ratings of 5 and 1 on the rating scale.

Six research assistants from the Institute for Research on Learning Disabilities were trained in the use of the Structure of Instruction Rating Scale and then allowed to practice rating the first tape of the "competent" teacher. After viewing the first tape, the observers and trainers went over the results and attempted to come to agreement on the "appropriate" rating for each variable on that tape. In order to reinforce this agreement they then rated the second tape.

In their description of a method for determining inter-rater agreement, Lawlis and Lu (1972) stated that agreement can be defined as exact, within one or two points. Tinsley and Weiss (1975) qualified this by pointing out that the decision of what standard to use as agreement and the rationale for this decision must be made prior to collecting data. In the current study, within one scale point was chosen as the level for agreement. Exact agreement on many dimensions of the SIRS would require observers, in many cases, to agree on distinctions that are very fine (sometimes varying no more than 5%). Some precision is sacrificed in not making this distinction, but such precision may be of questionable value. In the long run, it is probably more important to differentiate between high and low structured instructional programs than between high and very high structured programs, or between low and very low structured programs.

Counting ratings within one scale point as agreement, interobserver agreement on the original 10 variables ranged from .70 to 1.00, with a mean interrater agreement of .84. Feedback and discussion following the first tape resulted in improved observer

reliability on the second tape, with agreements ranging from .80 to 1.00, with a mean of .91. Mean ratings for the 10 variables on the two tapes are included in Table 1.

Insert Table 1 about here

The SIRS was used in four locations in Minnesota, and one in New York City. Teachers and administrators in some of the school districts were trained as observers since research assistants were not able to collect data in all locations. The same videotapes and training procedure were employed; the New York raters and one of the four Minnesota groups reported observer reliabilities. The raters trained in the Minnesota district reached an interrater agreement of .95 among themselves and .92 with the ratings of the six original observers. The New York observers achieved a mean interrater agreement of .85 among themselves and .84 with the ratings of the six research assistants. The somewhat lower agreement figures may be due to different trainers, or they may be simply a consequence of different location.

As O'Leary (1975) noted, one of the problems with measures of observer agreement is that of observer drift. Although observers may reach a high initial level of agreement, that agreement tends to disintegrate over time as each observer, or group of observers, develops idiosyncratic definitions. Thus, as recommended by O'Leary and others (Keller, 1980), checks on interobserver agreement were made four months after the initial training. The checks involved four of

the original six observers and took place in actual classroom settings. Checks on six students resulted in interrater agreements for one pair of observers ranging from .75 to .83 (X = .81), and for the other pair ranging from .92 to 1.00 (X = .95). In addition, training of additional observers (all research assistants) later in the year allowed for "recalibration" of the original observers with the training tape.

Subjects

The subjects for all statistical analyses were 158 grade 1-8 resource room students in four rural and suburban Minnesota school districts. All subjects were participants in research on the effectiveness of frequent data-based reading measures. The data were collected three times during the school year: fall (October), winter (February), and spring (May).

Results

Reliability

Interrater agreement is only one facet of reliability. Any measurement instrument also must show evidence of internal consistency or homogeneity; that is, to what degree do all items of the scale measure the same thing? Cronbach's Alpha, a measure of internal consistency, was chosen to assess the reliability of the SIRS.

Data were analyzed for each of the three data collections. For the fall data, item-total correlations resulted in an alpha of .84 over all 10 variables. The ratings on the 12 variables in winter showed somewhat lower inter-item correlations, with an alpha of .76. The spring data resulted in a reliability estimate of .81.

As can be seen in Table 2, although the item-total correlations for all variables (10 in the fall, 12 in winter and spring) generally were high, there was some instability among the correlations. Particularly, the variables Positive Consequences, Independent Practice, and Silent Practice showed low and inconsistent item-total correlations across the data collections. These data are consistent with the factor analysis that resulted in the omission of these latter three items from the scale.

Insert Table 2 about here

Dropping the three items from the scale resulted in stronger and more consistent item-total correlations, and hence a more reliable scale. As shown in Table 3, the remaining nine variables exhibited good to excellent inter-item correlations, resulting in alphas of .92, .82 and .88 for fall, winter, and spring, respectively. Thus, the SIRS evidences good reliability, consistent across items and over time.

Insert Table 3 about here

Factor Analysis

In order to help determine whether the 12 variables of the SIRS measured the same construct or several constructs, a factor analysis was performed on the data from the second data collection (N=157). The method chosen, principal-component analysis, makes no assumption

about the underlying structure of the variables, but simply identifies the best linear combination of variables, that is, the particular combination of variables that accounts for more of the variance in the data than any other linear combination of variables. The first principal component identified represents the single best combination of variables for this purpose, the second component the second best combination that is orthogonal (uncorrelated) to the first, and so on. The final solution sought in factor analysis allows one to determine whether there is an underlying pattern of relationships among the variables such that the data may be rearranged or reduced to a smaller number of factors or components accounting for these interrelations (Kim, 1975).

Initial factor analysis of the SIRS revealed four factors with eigenvalues greater than or equal to 1.0--that is, factors that accounted for at least the amount of the total variance of a single variable. Of these factors, the first accounted for 54.8% of the variance, the second for 24.7%. Since these two factors alone accounted for almost 80% of the variability in the data, a subsequent factor analysis examined only these two factors, in order to more clearly identify those variables loading on each factor.

The second factor analysis resulted in all of the variables but three loading on the first factor (see Table 4). Two of these three, Independent Practice and Positive Consequences, loaded on the second factor; for only Silent Practice was the variance unexplained by either factor. The first factor accountcd for 70.3% of the variance, and the second factor for only 29.7% of the variance.

Insert Table 4 about here

The existence of a single common factor shared by 9 of the 12 variables implies that they were measuring a single construct. Since this shared component accounted for 70.3% of the variance, all subsequent analyses treated these variables as one factor; the remaining three variables were analyzed separately.

Teacher Use of Structure in Instruction

Once sufficient reliability and evidence of consistent factor structure had been established, an effort was made to determine which of the variables being measured were experienced most by students during their instructional programs. Among the nine variables analyzed as one factor, the moderate to high ratings at all three points in time indicated that these aspects of classroom structure are relatively well established in most resource rooms (see Table 5). Standard deviations tended to be very stable over time.

Insert Table 5 about here

As is shown in Table 5, Corrections, Frequency of Correct Answers, Teacher Directed Learning, and Active Academic Responding were experienced most often in students' programs. At time 2, over 80% of the student programs received a rating of 4 or higher on the variables Corrections, Frequency of Correct Answers, and Active Academic Responding. It is apparent, however, that the elements of

structure represented by Independent Practice, Positive Consequences, and Silent Practice were much less a part of the programs studied. For example, at time 2, only 26% of the students were observed participating in any form of token economy as part of their programs.

For the most part, program structure remained very stable over time. There was an average difference of only .2 rating points between the highest and lowest ratings on any given variable over all three data collections.

Not surprisingly, there were great differences in the amount of structure students received in the various resource rooms. The mean SIRS rating for each classroom was obtained by averaging the SIRS rating for the individual students in that resource room. At time 3, these classroom means ranged from 2.89 (over 3 students) to 4.81 (over 6 students) and were, in fact, significantly different, $F(34) = 4.63$, $p < .001$. There were no significant differences in structure of instruction as a function of curriculum, nor were there significant grade level differences in structure of instruction provided to students.

Discussion

Few measurement devices in psychology have been used as widely, or maligned as widely, as rating scales. Despite the frequency of their use, and the voluminous research on their psychometric qualities (cf. Saal, Downey, & Lahey, 1980), the problems identified in the early literature (Guilford, 1954; Thorndike, 1920) have not yet been laid to rest. Thus, extreme care must be taken in the development of a new rating scale to ensure the reliability and validity of the

instrument.

The Structure of Instruction Rating Scale (SIRS) was designed to measure the instructional structure provided to individual students in the resource room. The dimensions included in the scale were chosen from the literature as those variables that seemed most likely to predict achievement. These variables then formed the behavioral definitions for the 12 items of the original rating scale.

Of prime importance in assuring the reliability of any observational instrument is achieving inter-observer agreement. If the scale does not evidence a high degree of inter-rater reliability, there is no assurance that the observers are measuring the same thing (Sulzer-Azaroff & Mayer, 1977). Raters were trained in the use of the SIRS through observation of videotapes and reached good to excellent levels of agreement during training. In addition, periodic checks ensured the maintenance of such agreement over time.

Additional statistical analyses confirmed the reliability and validity of the scale. Internal consistency estimates for the SIRS showed intercorrelations among the 12 dimensions high enough to prove scale reliability, yet not so high as to evidence halo effects (Keaveny & McGann, 1975), that is, the tendency of a rater to rate a subject in a similar manner on all dimensions. The dimensions represented by the SIRS variables were selected to measure a single construct, the structure of an instructional program; thus, factor analysis of the scale that identified a single construct accounting for the majority of the variance supports the construct validity of the SIRS. Other evidence supporting the construct validity of the

SIRS was its ability to differentiate between the amount of structure provided in different classrooms, and to predict achievement as indexed by both traditional and alternative measures (Skiba, King, & Deno, 1983).

Throughout the data analysis, 9 of the 12 variables tended to form one construct, while three dimensions (Independent Practice, Positive Consequences, and Silent Practice) consistently failed to group with the other nine. The failure of these three variables to relate to the other variables of the scale is interesting both from a psychometric and a research viewpoint. All three variables were observed infrequently in students' programs and, in fact, the circumstances of the observations precluded observation of Independent Practice in a majority or cases (54%). However, the variable Positive Consequences may have failed to predict academic achievement or load on the structure factor due to the lack of its use in the classrooms. This finding provides important and disturbing information about the state of the art of token economies in resource classrooms, especially given extensive research findings that indicate the effectiveness of such systems for special populations (Kazdin, 1977; O'Leary, Drabman, & Kass, 1973; Phillips, Phillips, Fixsen & Wolf, 1971; Staats & Butterfield, 1965).

The failure of Silent Practice to load on either factor is more difficult to explain, especially given the importance attached to silent reading practice in the literature (Graden, Thurlow, Ysseldyke, & Algozzine, 1982). This failure may represent a measurement problem. Most of the variables are free to vary independent of each other; that

is, the occurrence of a behavior included in one variable does not preclude the observation of another SIRS variable. Silent Practice and Oral Practice, however, are mutually exclusive: if a student is reading aloud, he/she cannot be observed reading silently, and vice versa. This may have reduced the amount of time available for observation of the variable Silent Practice.

Future researchers using the SIRS might wish to examine the relative contributions of the Oral Practice and Silent Practice variables. That is, ratings should be made on each variable independently, not concurrently. Such a procedure should eliminate the possibility of confounding due to mutually exclusive operational definitions.

A similar procedure also could be employed to provide increased opportunities to observe Independent Practice. In addition, future research might address the contribution of teacher praise as a possible variable in program structure and its relationship to student achievement. Anecdotal data obtained during this study suggested that the students often received praise from their teachers as part of their instructional program. However, as previously indicated, students rarely participated in a token ecoromy during their instructional lessons. By adding teacher praise as a variable, in addition to student participation in token economies, researchers should be able to ferret out the relative contribution of each dimension of these program variables in terms of their relationships with program structure and subsequent student achievement.

The SIRS has been shown to be a useful research instrument from

the standpoint of technical adequacy and heuristics. At the same time, the interest shown by school administrators and school psychologists during data collection implies that those who are responsible for assessing classroom environments and for teaching practices also may find the instrument to be useful. Certainly, the construct of program structure as discussed in this paper has implications for staff development with respect to changing teaching practices. The construct of structure and the implications for its relationship to staff development have been aptly summarized by Miller (1980):

> It requires a major restructuring of the classroom and a major reassessment of the teacher role, a role that most practitioners have worked long and hard to define for themselves. This raises staff development to another level of difficulty. When presented with such a scenario, teachers will need a great deal of support to resist the temptation to retreat into the familiarity of old routines. (p. 164)

Clearly, the construct of program structure and its relationship to time and student achievement continues to emerge as a validated area of important research. The research findings discussed here support, validate, and extend several previous research findings. Ultimately, the utility of these findings will be determined by school administrators who are willing to promote staff development to foster effective teaching practices.

References

Becker, W. C., Engelmann, S., & Thomas, D. R. Teaching: A course in applied psychology. Chicago: Science Research Associates, 1973.

Borg, W. Time and school learning. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, D.C.: National Institute of Education, 1980.

Brophy, J. Teacher praise: A functional analysis. Review of Educational Research, 1981, 51, 5-32.

Cooley, W., & Leinhardt, G. The instructional dimensions study. Educational Evaluation and Policy Analysis, 1980, 2, 7-24.

Fisher, C., Berliner, D., Filby, N., Marliave, R., Cahen, L., & Dishaw, M. Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, D.C.: National Institute of Education, 1980.

Gersten, R. M. Direct instruction programs in special education settings: A review of evaluation research findings. Paper presented at the annual meeting of the Council for Exceptional Children, New York, 1981.

Gersten, R., Carnine, D., Zoref, L., & Cronin, D. A study of implementation of educational change in an urban setting: Integrating teacher effectiveness and implementation research. Unpublished manuscript, University of Oregon, Eugene, 1981.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1982.

Guilford, J. P. Psychometric methods. New York: McGraw-Hill, 1954.

Kazdin, A. E. The token economy: A review and evaluation. New York: Plenum, 1977.

Keaveny, T. J., & McGann, A. F. A comparison of behavioral expectation scales and graphic rating scales. Journal of Applied Psychology, 1975, 60, 695-703.

Keller, H. R. Issues in the use of observational assessment. School Psychology Review, 1980, 9(1), 21-30.

Kim, J. Factor analysis. In N. H. Nie, C. H. Hull, J. G. Jenking, K. Steinbrenner, & D. H. Bents (Eds.), Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1975.

Lawlis, G. F., & Lu, F. Judgment of counseling process: Reliability, agreement, and error. Psychological Bulletin, 1972, 78, 17-20.

Miller, L. BTES: Implications for staff development. In C. Denham, A. Lieberman (Eds)., Time to learn. Washington, D.C.: National Institute of Education, 1980.

O'Leary, K. D. Behavioral assessment: An observational slant. In R. Weinberg & F. Wood (Eds.), Observation of pupils and teachers in mainstream and special education settings: Alternative strategies. Minneapolis: Leadership Training Institute, 1975.

O'Leary, K. D., Drabman, R., & Kass, R. E. Maintenance of appropriate behavior in a token program. Journal of Abnormal Child Psychology, 1973, 1, 127-138.

Phillips, E. L., Phillips, E. A., Fixsen, D. L., & Wolf, M. M. Achievement place: Modification of the behaviors of pre-delinquent boys within a token economy. Journal of Applied Analysis, 1971, 4, 45-59.

Saal, F. E., Downey, R. G., & Lahey, M. A. Rating the ratings: Assessing the psychometric quality of rating data. Psychological Bulletin, 1980, 88(2), 413-428.

Skiba, R., King, R., & Deno, S. L. The relationship of structure of instruction to achievement (Research Report, in preparation. Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1983.

Staats, A. W., & Butterfield, W. H. Treatment of nonreading in a culturally deprived juvenile delinquent: An application of reinforcement principles. Child Development, 1965, 36, 925-942.

Starlin, C. M. Evaluating and teaching reading to "irregular" kids. Iowa Perspective, December, 1979.

Stevens, R., & Rosenshine, B. Advances in research on teaching. Exceptional Education Quarterly, 1981, 2(1), 1-10.

Sulzer-Azaroff, B., & Mayer, G. R. Applying behavior analysis procedures with children and youth. New York: Holt, Rinehart, & Winston, 1977.

Thorndike, E. L. A constant error in psychological ratings. Journal of Applied Psychology, 1920, 4, 25-29.

Tinsley, H. E., & Weiss, D. J. Interrater reliability and agreement of subjective judgments. Journal of Counseling Psychology, 1975,

Zoref, L. S. A prototype for how to evaluate implementation of a structured educational program. Unpublished doctoral dissertation, University of Oregon, 1981.

Table 1

Mean Ratings for the Two Videotapes[a]

	Tape 1 ("Exemplary" Teacher)	Tape 2 ("Poor" Teacher)
Instructional Grouping	3.20	1.00
Teacher Directed Learning	5.00	2.00
Active Academic Responding	4.40	2.17
Demonstration and Prompting	4.60	1.00
Controlled Practice	4.80	1.33
Frequency of Correct Answers	4.40	2.17
Independent Practice	5.00	1.00
Corrections	5.00	1.17
Positive Consequences	4.40	1.00
Pacing	4.60	1.33
Overall $\bar{X}$	4.54	1.42

[a]Means are based upon the ratings of the six IRLD research assistants used in the initial SIRS training.

Table 2

Corrected Item - Total Correlations for All Twelve Variables

	Fall (N=23)	Winter (N=59)	Spring (N=64)
Instructional Grouping	.52	.51	.32
Teacher Directed Learning	.67	.63	.78
Active Academic Responding	.41	.45	.66
Demonstration and Prompting	.92	.52	.74
Controlled Practice	.79	.63	.55
Frequency of Correct Answers	.51	.25	.42
Independent Practice	-.17	.17	.23
Corrections	.65	.26	.64
Positive Consequences	.34	.30	.08
Pacing	.80	.73	.82
Oral Practice	-	.42	.53
Silent Practice	-	.05	.12
Coefficient Alpha	.84	.76	.81

Table 3

Corrected Item - Total Correlations Dropping Three Variables

	Fall (N=59)	Winter (N=59)	Spring (N=64)
Instructional Grouping	.69	.51	.41
Teacher-Directed Learning	.71	.66	.79
Active Academic Responding	.64	.44	.66
Demonstration and Prompting	.91	.51	.75
Controlled Practice	.73	.60	.54
Frequency of Correct Answers	.66	.33	.47
Corrections	.67	.39	.67
Pacing	.84	.70	.81
Oral Practice	-	.55	.61
Coefficient Alpha	.92	.82	.88

Table 4

Results of Second Factor Analysis of the Twelve SIRS Variables[a]

Variable	Factor 1	Factor 2
Instructional Grouping	.449	.286
Teacher-Directed Learning	.665	-.129
Active Academic Responding	.587	-.079
Demonstration and Prompting	.663	-.034
Controlled Practice	.696	.280
Frequency of Correct Answers[b]	.362	-.378
Independent Practice	.206	.796
Corrections	.552	-.444
Positive Consequences	.156	.390
Pacing	.640	-.001
Oral Practice	.520	-.032
Silent Practice	-.121	-.045

[a]Underlined variables load most heavily on the factor.

[b]Although this variable loaded on both factors (and in fact loaded more heavily on Factor 2) the decision was made to include it with the Factor with which it was positively correlated.

Table 5

Means for All Subjects on SIRS Variables

	Fall $\bar{X}$ (SD)	Winter $\bar{X}$ (SD)	Spring $\bar{X}$ (SD)	Average $\bar{X}$
Corrections	4.07 (1.17)	4.30 (.92)	4.22 (1.10)	4.20
Frequency of Correct Answers	4.03 (.86)	4.36 (.78)	4.08 (.95)	4.16
Teacher-Directed Learning	4.03 (1.07)	4.11 (.93)	4.06 (.95)	4.07
Active Academic Responding	3.93 (1.07)	4.19 (.99)	4.04 (1.01)	4.05
Pacing	3.58 (1.25)	3.78 (1.07)	3.84 (1.07)	3.73
Demonstration and Prompting	3.45 (1.27)	3.81 (1.19)	3.84 (1.16)	3.70
Controlled Practice	3.65 (1.28)	3.60 (1.27)	3.71 (1.17)	3.65
Instructional Grouping	3.38 (1.51)	3.46 (1.32)	3.70 (1.26)	3.51
Oral Practice	—	3.00 (1.28)	3.06 (1.28)	3.03
SIRS Mean	3.77	3.85	3.84	
Independent Practice	2.28 (1.19)	2.11 (1.31)	2.12 (1.25)	2.17
Positive Consequences	1.89 (1.29)	1.80 (1.39)	1.72 (1.30)	1.80
Silent Practice	—	2.11 (1.26)	2.19 (1.18)	1.43

Appendix A

Structure of Instruction Rating Scale (SIRS)

School:_____ Student:_____

Date:_____ Teacher:_____

Observer:_____ Number of Students in Group:_____

Number of observations prior to rating:_____

Time observation begins:_____ Time observation ends:_____

Time allocated to reading instruction per day:_____

Currirulum used for instruction: Publisher_____

Series_____ Level_____

Instructions

Circle the number that accurately reflects your rating for each variable. Only one number may be circled per variable. If you are unable to evaluate a certain variable, mark N/A (not applicable) next to the left-hand column.

1.	Instructional Grouping	1	2	3	4	5
2.	Teacher-directed Learning	1	2	3	4	5
3.	Active Academic Responding	1	2	3	4	5
4.	Demonstration/Prompting	1	2	3	4	5
5.	Controlled Practice	1	2	3	4	5
6.	Frequency of Correct Answers	1	2	3	4	5
7.	Independent Practice	1	2	3	4	5
8.	Corrections	1	2	3	4	5
9.	Positive Consequences	1	2	3	4	5
10.	Pacing	1	2	3	4	5
11.	Oral Practice on Outcome Behavior	1	2	3	4	5
12.	Silent Practice on Outcome Behavior	1	2	3	4	5

SIRS

Operational Definitions Codebook

1. Instructional Grouping

   5 - 90% or more of the instruction this student receives from the teacher is on an individual basis.

   1 - 10% or less of the instruction this student receives from the teacher is on an individual basis.

2. Teacher-Directed Learning

   5 - Student's instruction is extremely organized, businesslike, and teacher is firm in direction and control of activities. For example, student is presented with questions, student has material to cover, etc.

   1 - Student's instruction is casually organized and very spontaneous. Teacher is not committed to having the student work on a particular set of material. Instructional materials do not determine what activities student engages in and the lessons change according to problems or mood of this student.

3. Active Academic Responding

   5 - The student is actively practicing the academic skills to be learned more than 75% of the time observed. Specifically, the student is engaged in oral or written responding to teacher questions or written material, e.g., reading aloud, answering questions, writing, or computing. Student rarely is involved in non-academic conversations with teacher or other students Attending to the lesson without responding, such as sitting, looking, listening, and/or following along in a book does not apply. The student must make an active, written or oral response.

   1 - The student is actively practicing the skills to be learned less than 10% of the time observed. Instructional lessons may be interrupted or shortened to include "process" and other non-academic activities, e.g., clarifying feelings, opinions, and working on arts and crafts.

4. Demonstration and Prompting

   5 - Appropriate steps of the desired behavior to be performed are demonstrated for the student. Student is given an opportunity to practice the step(s) as teacher provides prompts for correct behavior that approximates or achives desired response.

   1 - Teacher attempts to teach the student a behavior without using demonstration and prompting techniques.

SIRS

5. Controlled Practice

5 - Student's practice of material is actively controlled by teacher who frequently asks questions to clarify that the student understands what has just been demonstrated. Questions are convergent (single factual answer) and the student's answers consistently follow the questions and are given teacher feedback.

1 - Student is rarely questioned by teacher following demonstration of new materials. Questions are more divergent (open-ended, several interpretations) than convergent (single factual answer). Student's response is not consistently followed by teacher feedback. The type of questions are such that several answers are acceptable, i.e., questions are abstract or ambiguous.

Examples:

If during an oral reading session:

a) the teacher frequently attempts to clarify the material with convergent questions ("what color hat was John wearing?"), a 5 would be recorded.

b) the teacher asks few questions, most of which are divergent ("What do you think this means?"), a 1 would be recorded.

c) the teacher asks few convergent questions or many divergent questions, the appropriate rating would be a 3.

6. Frequency of Correct Answers

5 - Academic lessons are conducted in such a way that the difficulty of the material allows the student to achieve mean accuracy of 80% or higher.

1 - Academic material is difficult for student, component steps are large or unsequenced, and mean accuracy for student is less than 55%.

(Note: If the student has no opportunity for oral or written response during the observational period, item 6 would be rated N/A - not applicable, while items 3 and 5 would most likely be rated 1.).

7. Independent Practice

5 - When engaged in independent seatwork, the student frequently is monitored by the teacher who assists, clarifies, and praises the student for academic engaged tasks.

(Note: Independent seatwork is defined here as a student working on an assigned task for at least 5 minutes. [If no such 5-minute block of time is observed, Item 7 is rated N/A].)

SIRS

1 - When student is engaged in academic seat-work activities, little attention is given by teacher who directs seat-work activities from a distance or engages in work separate from the assigned seat work. Teacher is generally not helpful or supportive to student during independent practice time.

8. Corrections

5 - The student's errors are consistently corrected by the teacher. When the student either does not respond, responds incorrectly, or does not respond in unison if the activity is group directed and requires such responding, the teacher will systematically attempt to correct the student by asking a simpler question, refocusing student's attention to elicit correct response from the student or provide general rules by which to determine the correct answer 90% or more of the time.

1 - Student's errors are rarely and inconsistently corrected by the teacher. The student responses are not systematically corrected. Student's errors are corrected 50% or less of the time.

For example: In oral reading this includes teacher correction of skips and mispronunciations, or help in sounding out hesitations.

9. Positive Consequences

5 - Positive events (tokens, points, activities, etc.) are given to the student when performing the desired behavior. When learning a new skill the student receives positive consequence for approximations of the desired behavior. Consequences are consistently received during academic training time. Praise and compliments, e.g., "good working, nice job," are not included in this definition.

1 - Student rarely receives positive consequences for academic work. When student receives consequences they usually are for social behavior, rather than for behaviors occurring under systematic academic training.

10. Pacing

5 - The pace of the lesson is rapid, providing many opportunities for response by the student. As a result, attention is high and off-task behavior is low.

1 - The pace of the lesson is slow and the student's rate of responding is low. Lesson format frequently varies, is not highly structured, and student attention may be low.

11. Oral Practice on Outcome Behavior

5 - Student reads aloud from context nearly all the time (85-100% or 12-15 min. of a 15 min. observation).

1 - Student does not read aloud during the observation (0% of the time).

(Note: Reading aloud for measurement purposes should not be considered when rating this variable. Reading in context is defined as reading phrases, sentences, paragraphs, or story selections.)

Examples:

If the student is reading isolated words nearly the entire time, the appropriate rating is a 3.

If the student is reading aloud from a text about half the time, a 3 would be recorded.

12. Silent Practice on Outcome Behavior

5 - Student reads silently from context nearly all the time (85-100% or 12-15 min. of a 15 min. observation).

1 - Student does not read silently during the observation (0% of the time).

(Note: Reading in context is defined as the same as #11. The examples of #11 are the same for #12, with silent reading.)

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota roundtable conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Mirkin, P. K., Chiang, B., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Mirkin, P. K., Lowry, L., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Mirkin, P. K., & Marston, D. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Mirkin, P. K., Robinson, S., & Evans, P. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Instructional planning: Information collected by school psychologists vs. information considered useful by teachers (Research Report No. 30). June, 1980.

Algozzine, B., Webber, J., Campbell, M., Moore, S., & Gilliam, J. Classroom decision making as a function of diagnostic labels and perceived competence (Research Report No. 31). June, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., Richey, L., & Thurlow, M. L. Psychoeducational assessment and decision making: A computer-simulated investigation (Research Report No. 32). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., & Richey, L. Psychoeducational assessment and decision making: Individual case studies (Research Report No. 33). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R., Potter, M., & Richey, L. Technical supplement for computer-simulated investigations of the psychoeducational assessment and decision-making process (Research Report No. 34). July, 1980.

Algozzine, B., Stevens, L., Costello, C., Beattie, J., & Schmid, R. Classroom perspectives of LD and other special education teachers (Research Report No. 35). July, 1980.

Algozzine, B., Siders, J., Siders, J., & Beattie, J. Using assessment information to plan reading instructional programs: Error analysis and word attack skills (Monograph No. 14). July, 1980.

Ysseldyke, J., Shinn, M., & Epps, S. A comparison of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability (Research Report No. 36). July, 1980.

Algozzine, B., & Ysseldyke, J. E. An analysis of difference score reliabilities on three measures with a sample of low achieving youngsters (Research Report No. 37). August, 1980.

Shinn, M., Algozzine, B., Marston, D., & Ysseldyke, J. A theoretical analysis of the performance of learning disabled students on the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 38). August, 1980.

Richey, L. S., Ysseldyke, J., Potter, M., Regan, R. R., & Greener, J. Teachers' attitudes and expectations for siblings of learning disabled children (Research Report No. 39). August, 1980.

Ysseldyke, J. E., Algozzine, B., & Thurlow, M. L. (Eds.). A naturalistic investigation of special education team meetings (Research Report No. 40). August, 1980.

Meyers, B., Meyers, J., & Deno, S. Formative evaluation and teacher decision making: A follow-up investigation (Research Report No. 41). September, 1980.

Fuchs, D., Garwick, D. R., Featherstone, N., & Fuchs, L. S. On the determinants and prediction of handicapped children's differential test performance with familiar and unfamiliar examiners (Research Report No. 42). September, 1980.

Algozzine, B., & Stoller, L. Effects of labels and competence on teachers' attributions for a student (Research Report No. 43). September, 1980.

Ysseldyke, J. E., & Thurlow, M. L. (Eds.). The special education assessment and decision-making process: Seven case studies (Research Report No. 44). September, 1980.

Ysseldyke, J. E., Algozzine, B., Potter, M., & Regan, R. A descriptive study of students enrolled in a program for the severely learning disabled (Research Report No. 45). September, 1980.

Marston, D. Analysis of subtest scatter on the tests of cognitive ability from the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 46). October, 1980.

Algozzine, B., Ysseldyke, J. E., & Shinn, M. Identifying children with learning disabilities: When is a discrepancy severe? (Research Report No. 47). November, 1980.

Fuchs, L., Tindal, J., & Deno, S. Effects of varying item domain and sample duration on technical characteristics of daily measures in reading (Research Report No. 48). January, 1981.

Marston, D., Lowry, L., Deno, S., & Mirkin, P. An analysis of learning trends in simple measures of reading, spelling, and written expression: A longitudinal study (Research Report No. 49). January, 1981.

Marston, D., & Deno, S. The reliability of simple, direct measures of written expression (Research Report No. 50). January, 1981.

Epps, S., McGue, M., & Ysseldyke, J. E. Inter-judge agreement in classifying students as learning disabled (Research Report No. 51). February, 1981.

Epps, S., Ysseldyke, J. E., & McGue, M. Differentiating LD and non-LD students: "I know one when I see one" (Research Report No. 52). March, 1981.

Evans, P. R., & Peham, M. A. S. Testing and measurement in occupational therapy. A review of current practice with special emphasis on the Southern California Sensory Integration Tests (Monograph No. 15). April, 1981.

Fuchs, L., Wesson, C., Tindal, G., & Mirkin, P. Teacher efficiency in continuous evaluation of IEP goals (Research Report No. 53). June, 1981.

Fuchs, D., Featherstone, N., Garwick, D. R., & Fuchs, L. S. The importance of situational factors and task demands to handicapped children's test performance (Research Report No. 54). June, 1981.

Tindal, G., & Deno, S. L. Daily measurement of reading: Effects of varying the size of the item pool (Research Report No. 55). July, 1981.

Fuchs, L. S., & Deno, S. L. A comparison of teacher judgment, standardized tests, and curriculum-based approaches to reading placement (Research Report No. 56). August, 1981.

Fuchs, L., & Deno, S. The relationship between curriculum-based mastery measures and standardized achievement tests in reading (Research Report No. 57). August, 1981.

Christenson, S., Graden, J., Potter, M., & Ysseldyke, J. Current research on psychoeducational assessment and decision making: Implications for training and practice (Monograph No. 16). September, 1981.

Christenson, S., Ysseldyke, J., & Algozzine, B. Institutional constraints and external pressures influencing referral decisions (Research Report No. 58). October, 1981.

Fuchs, L., Fuchs, D., & Deno, S. Reliability and validity of curriculum-based informal reading inventories (Research Report No. 59). October, 1981.

Algozzine, B., Christenson, S., & Ysseldyke, J. Probabilities associated with the referral-to-placement process (Research Report No. 60). November, 1981.

Tindal, G., Fuchs, L. Christenson, S., Mirkin, P., & Deno, S. The relationship between student achievement and teacher assessment of short- or long-term goals (Research Report No. 61). November, 1981.

Mirkin, P., Fuchs, L., Tindal, G., Christenson, S., & Deno, S. The effect of IEP monitoring strategies on teacher behavior (Research Report No. 62). December, 1981.

Wesson, C., Mirkin, P., & Deno, S. Teachers' use of self instructional materials for learning procedures for developing and monitoring progress on IEP goals (Research Report No. 63). January, 1982.

Fuchs, L., Wesson, C., Tindal, G., Mirkin, P., & Deno, S. Instructional changes, student performance, and teacher preferences: The effects of specific measurement and evaluation procedures (Research Report No. 64). January, 1982.

Potter, M., & Mirkin, P. Instructional planning and implementation practices of elementary and secondary resource room teachers: Is there a difference? (Research Report No. 65). January, 1982.

Thurlow, M. L., & Ysseldyke, J. E. Teachers' beliefs about LD students (Research Report No. 66). January, 1982.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). January, 1982.

King, R., Wesson, C., & Deno, S. Direct and frequent measurement of student performance: Does it take too much time? (Research Report No. 67). February, 1982.

Greener, J. W., & Thurlow, M. L. Teacher opinions about professional education training programs (Research Report No. 68). March, 1982.

Algozzine, B., & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept (Research Report No. 69). March, 1982.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions (Research Report No. 70). March, 1982.

Shinn, M. R., Ysseldyke, J., Deno, S., & Tindal, G. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving (Research Report No. 71). March, 1982.

Thurlow, M. L. Graden, J., Greener, J. W., & Ysseldyke, J. E. Academic responding time for LD and non-LD students (Research Report No. 72). April, 1982.

Graden, J., Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student (Research Report No. 74). April, 1982.

Wesson, C., Deno, S., & Mirkin, P. Research on developing and monitoring progress on IEP goals: Current findings and implications for practice (Monograph No. 18). April, 1982.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students (Research Report No. 75). May, 1982.

Algozzine, B., Ysseldyke, J., Christenson, S., & Thurlow, M. Teachers' intervention choices for children exhibiting different behaviors in school (Research Report No. 76). June, 1982.

Tucker, J., Stevens, L. J., & Ysseldyke, J. E. Learning disabilities: The experts speak out (Research Report No 77). June, 1982.

Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). July, 1982.

Mirkin, P. K., & Potter, M. L. A survey of program planning and implementation practices of LD teachers (Research Report No. 80). July, 1982.

Fuchs, L. S., Fuchs, D., & Warren, L. M. Special education practice in evaluating student progress toward goals (Research Report No. 81). July, 1982.

Kuehnle, K., Deno, S. L., & Mirkin, P. K. Behavioral measurement of social adjustment: What behaviors? What setting? (Research Report No. 82). July, 1982.

Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). July, 1982.

Videen, J., Deno, S., & Marston, D. Correct word sequences: A valid indicator of proficiency in written expression (Research Report No. 84). July, 1982.

Potter, M. L. Application of a decision theory model to eligibility and classification decisions in special education (Research Report No. 85). July, 1982.

Greener, J. E., Thurlow, M. L., Graden, J. L., & Ysseldyke, J. E. The educational environment and students' responding times as a function of students' teacher-perceived academic competence (Research Report No. 86). August, 1982.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). August, 1982.

Skiba, R., Wesson, C., & Deno, S. L. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 88). September, 1982.

Marston, D., Tindal, G., & Deno, S. L. Eligibility for learning disability services: A direct and repeated measurement approach (Research Report No. 89). September, 1982.

Thurlow, M. L., Ysseldyke, J. E., & Graden, J. L. LD students' active academic responding in regular and resource classrooms (Research Report No. 90). September, 1982.

Ysseldyke, J. E., Christenson, S., Pianta, R., Thurlow, M. L., & Algozzine, B. An analysis of current practice in referring students for psychoeducational evaluation: Implications for change (Research Report No. 91). October, 1982.

Ysseldyke, J. E., Algozzine, B., & Epps, S. A logical and empirical analysis of current practices in classifying students as handicapped (Research Report No. 92). October, 1982.

Tindal, G., Marston, D., Deno, S. L., & Germann, G. Curriculum differences in direct repeated measures of reading (Research Report No. 93). October, 1982.

Fuchs, L.S., Deno, S. L., & Marston, D. Use of aggregation to improve the reliability of simple direct measures of academic performance (Research Report No. 94). October, 1982.

Ysseldyke, J. E., Thurlow, M. L., Mecklenburg, C., & Graden, J. Observed changes in instruction and student responding as a function of referral and special education placement (Research Report No. 95). October, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study (Research Report No. 96). November, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Report No. 97). November, 1982.

Tindal, G., Wesson, C., Germann, G., Deno, S. L., & Mirkin, P. K. The Pine County model for special education delivery: A data-based system (Monograph No. 19). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. An analysis of the conceptual framework underlying definitions of learning disabilities (Research Report No. 98). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. Public-policy implications of different definitions of learning disabilities (Research Report No. 99). November, 1982.

Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Wesson, C., Deno, S. L., & Algozzine, B. Generalizations from five years of research on assessment and decision making (Research Report No. 100). November, 1982.

Marston, D., & Deno, S. L. Measuring academic progress of students with learning difficulties: A comparison of the semi-logarithmic chart and equal interval graph paper (Research Report No. 101). November, 1982.

Beattie, S., Grise, P., & Algozzine, B. Effects of test modifications on minimum competency test performance of third grade learning disabled students (Research Report No. 102). December, 1982

Algozzine, B., Ysseldyke, J. E., & Christenson, S. An analysis of the incidence of special class placement: The masses are burgeoning (Research Report No. 103). December, 1982.

Marston, D., Tindal, G., & Deno, S. L. Predictive efficiency of direct, repeated measurement: An analysis of cost and accuracy in classification (Research Report No. 104). December, 1982.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. Teaching structure and student achievement effects of curriculum-based measurement: A causal (structural) analysis (Research Report No. 105). December, 1982.

Mirkin, P. K., Fuchs, L. S., & Deno, S. L. (Eds.). Considerations for designing a continuous evaluation system: An integrative review (Monograph No. 20). December, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). December, 1982.

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). January, 1983.